

Unfocussed spatial attention underlies the crowding effect in indirect form vision

Hans Strasburger

Generation Research Program, University of München, Germany
Dept. of Medical Psychology, University of Göttingen, Germany



Abstract

We studied mechanisms underlying the crowding effect in indirect form vision by measuring recognition contrast sensitivity of a character with flankers to the left and right. Attentional and featural contributions to the effect can be separated by a new paradigm that distinguishes pattern location errors from pattern recognition errors and further by manipulating the focusing of spatial attention through a positional cue, appearing 150 ms before the target. Measurements were on the horizontal meridian, at 1°, 2°, and 4° eccentricity, and a range of flankers' distances were used. Our results show that in normal indirect viewing, the impairment of character recognition by crowding is – in particular at intermediate flanker distances – caused to a large part by spatially imprecise focusing of attention. In contrast, the enhancement of performance by a transient positional cue seems mediated through a separate attentional mechanism such that attentional locus and focus are controlled independently. Our results furthermore lend psychophysical support to a separate coding of the *what* and *where* in pattern recognition.

Key words: Peripheral vision, attention, character recognition, contrast sensitivity.

Introduction

When targets are seen in the presence of neighbouring patterns or contours, their visibility is reduced. This phenomenon is known as the crowding effect in the ophthalmic and vision literature but is referred to as lateral masking in experimental psychology (Strasburger, Harvey, & Rentschler, 1991). The effect is small in foveal vision (Flom, Heath, & Takahashi, 1963; Flom, Weymouth, & Kahnemann, 1963) but plays an overriding role in the normal peripheral visual field (Bouma, 1970), in amblyopes and dyslectics (Stuart & Burian, 1962; Geiger & Lettvin, 1989), as well as in children and juveniles (Atkinson, Pimm-Smith, Evans, Harding, & Braddick, 1986).

Experimental psychology and vision research have independently of each other discovered what is probably the same effect. Averbach & Coriell (1961) had used George Sperling's (1960) iconic-memory paradigm using spatial cues – a “bar marker” pointing to the target character or a “circle indicator” around it – to attract attention within a 2×8 letter string. Whereas the bar cue was effective, the circle unexpectedly decreased rather than enhanced recognition performance and thus lateral masking was shown. Eriksen and co-workers, in a series of papers, followed up on this work using the bar cue and started modern research on spatial attention (Strasburger, 2003b). In particular, Eriksen & Rohrbaugh (1970) introduced an analysis of perceptual errors (also used in this study) which would have allowed to separate sensory and attentional influences on the lateral-masking effect. How-

ever, Eriksen ruled out the relation of their findings to crowding on the wrong assumption that the stimulus positions they used was essentially equivalent to foveal presentation (they were at 2.2° eccentricity), such that the results by Flom et al. (1963; 1963) would apply: Flom et al. had shown foveal crowding to have a much smaller range of interaction than the distances in Eriksen's stimulus arrangement. It was only later that Bouma (1970) showed the strong influence of retinal eccentricity on the crowding effect. Thus, even though the crowding effect, described as early as 1936 by Ehlers, was well studied in amblyopic vision in 1963 and in normal foveal vision in 1970 – where it is small and of little practical significance – its important role in indirect and peripheral vision became apparent only much later. In the following, both in the cognitive sciences (for reviews see e.g. A. H. C van der Heijden, 1987; Desimone & Duncan, 1995; Gazzaniga, 1999, Chpt. 43 – 49; Schneider & Mojzisch, 2001) and in vision research, effects of spatial attention and of lateral interactions were widely assumed to be unrelated to each other, even though on different grounds.

In vision research, we were the first to show that the crowding effect in indirect vision can be understood at least partly as an effect of spatial attention (Strasburger et al., 1991; Strasburger & Rentschler, 1995). In that paper we analysed at an eccentric position (4°) the percentage of localization errors, i.e. the percentage, out of the failures of recognizing the center (target) character, of correspondences between the observer's response and one of the flanking characters. We showed that there was agreement in a large percentage of cases. This can be understood as pattern recognition accompanied by a loss of the posi-

tional information, for example from an inability to precisely focus spatial attention at eccentric visual field locations. The current paper secures and extends these findings. He et al. (1996) who used a grating-adaptation task took a similar view, stating that “this crowding effect reflects the limited resolution of the spatial attention mechanisms”. (For a review of research on the crowding effect and lateral masking see Strasburger, 2003b, Chpt. 6).

In the present report, three paradigms are used to analyse the influence of spatial attention on the crowding effect in indirect vision, all three operating at the contrast threshold of character recognition. A crowding condition similar to that in Strasburger et al. (1991) is compared (1) to a cued condition, with a circle around the target, and (2) a condition in which positional information is separated from the pattern content. Thirdly, Eriksen’s error analysis is used to demonstrate independent processing of pattern content and position by the visual system.

Methods

Contrast thresholds for the recognition of characters were obtained in a 10-alternative forced-choice (10-afc) adaptive procedure (R_Contrast, Strasburger, 1997) based on the ML-PEST maximum-likelihood algorithm (Harvey, 1986; 1997). The procedure is the same as that used earlier (Strasburger et al., 1991) except that the software was ported to current PC hardware (Jüttner & Strasburger, 1997) and uses dithering to achieve higher than 8-bit grey scale resolution on a standard graphics card (Bach, Meigen, & Strasburger, 1997). Thresholds are defined as the point of inflection on a Weibull function, which, for the 10-afc task, is at 67%-correct identification performance. Stimuli were the ten digits 0–9, presented for 100 ms as white patterns on the grey background of 50 cd/m² luminance (Figure 1). Their images were in multiples of a 5×7 pixel matrix and letter size was specified as letter height; so the width of a character is ~0.71 times the specified letter size. Stimulus contrast was varied from trial to trial as specified by ML-PEST; a run was ended when the estimated threshold confidence interval reached a pre-specified level of 0.2 log units (i.e., ±0.1); on average this took 30 trials (29.77). For a sample run see Strasburger (2001b, Fig. 3). Contrast is specified in the Michelson measure, i.e. $C_M = (L_s - L_b) / (L_s + L_b)$ (with L_s and L_b denoting the stimulus and background luminance, respectively). For readers interested in results in the Weber measure $C_W = (L_s - L_b) / L_b$, which is popular in character recognition, it is useful to know that the two measures are related by $C_W = 2 C_M / (1 - C_M)$.

Twenty young observers of both genders (age 20 – 30) were tested at identical conditions. A total of 34,000 trial responses were obtained (1.700 per subject).

Crowding Conditions

Three crowding conditions were compared to a single character condition: a flanked condition, a flanked-and-cued condition (“cued” in short), and a flanked condition termed “content-only threshold”, explained below. In each case, the three characters (target and flankers) were of same size and contrast, subjects are instructed to report the center character, and the dependent variable was the contrast threshold for recognizing the central, target character (Figure 1). In the flanked condition, the target is surrounded by a neighbouring character left and right. Distances were well above that where the characters would touch (which would be at 0.42° flanker distance for the 0.6° size). In the cued condition, there are again flanking characters left and right (of same size and contrast as the target) but additionally attention is attracted to the central character by a black circle around the latter, appearing 150 ms before the stimulus onset and disappearing with stimulus onset. The SOA of 150 ms was chosen based on Eriksen & Collins’ (1969) finding that at this delay the cueing effect is optimal. Nakayama & Mackeben (1989) similarly report that a transient cue which appears at 70 – 150 ms before the stimulus has an optimum cueing effect.

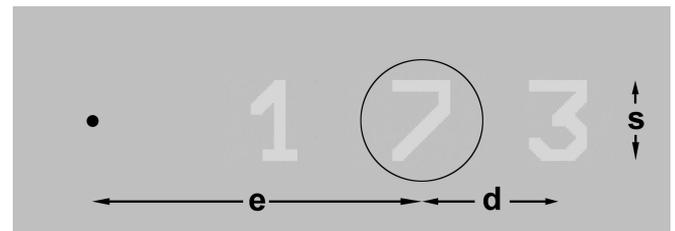


Figure 1. Stimulus layout in the flanked and cued crowding condition. Letter size (s) is specified as letter height in degrees of visual angle; flanking distances (d) are measured from the respective character centers. e : eccentricity.

Content-only contrast threshold: A third condition was designed to separate positional information from pattern content. The perceptual and attentional aspects were identical to the flanked condition, i.e. The target character was flanked left and right by a character of same size and contrast and, as before, subjects attended the central character which they were instructed to identify. However, the thresholding algorithm was altered in that correspondence of the observer’s response not only with the target but also with one of the two flanking characters were treated as correct responses by the adaptive algorithm. Note that the contrast of all three characters is thus lower. Both in this and the standard flanking condition the flanking characters were different from the central one and different from each other, to know which position the observer had responded to. Subjects were not aware of that since they attended to one character only and stimuli were often not recognizable. No feedback was given to the subject (in any condition) with respect to

whether the response was correct or incorrect. The resulting contrast threshold measures sensitivity to recognizing pattern content, irrespective of the pattern's location, in a flanked condition with sustained attention focussed at the center character. Relative to the standard flanked condition (1), there are thus two manipulations, one which (according to accepted models) manipulates attention (2), and one which leaves the percept unchanged but analyses the psychometric function in a different way (3).

The proportion of correct responses by chance alone (the guessing rate g) is different in the content-only task from that in a standard n-afc task, being $g = 1/10 + 1/9 + 1/8 = 33.6\%$. The thresholds are expected to be unaffected by that, however, since the maximum-likelihood algorithm converges to the point of inflection on the psychometric function (Harvey, 1986), which is independent of the guessing rate. The situation is illustrated in Figure 2 which shows psychometric functions (normalized to threshold) for a 10-afc and a 3-out-of-10 task, with the guessing rates set to 10% and 33.6%, respectively. The logistic function is chosen as the underlying sigmoid, so that the functions are given by

$$p_c = g + (1-g-I) / (1 + 10^{b \lambda})$$

The slope at the point of inflection has been set equal for the two functions, to $b' = 4.9 p_c / \log_{10}$, based on the value reported by Strasburger (2001b) for the 10 afc character-recognition task (the resulting values of b in the logistic function are $b = 9.67$ and $\beta = 13.21$; see Strasburger, 2001a for the conversions). The lapsing rate I (the rate of failures at high contrast) was set to 2%. The point of inflection, i.e. the point towards which the algorithm converges, occurs at slightly different criterion levels (54% and 65.8%, respectively), but is, independent of this, at threshold, indicated by the dashed line.

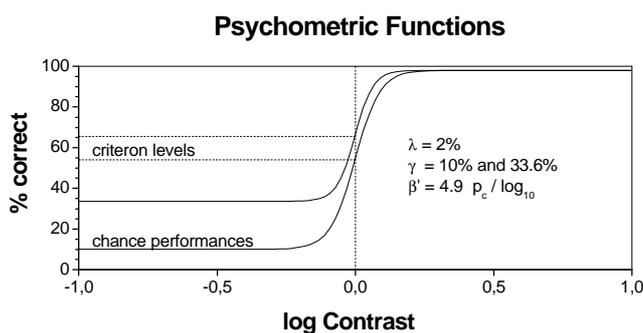


Figure 2. Psychometric functions for a 10-afc task and a 3-out-of-10 task. The guessing rate γ is 10% and 33.6%, respectively. The slope has been set equal in the two functions, to $4.9 p_c / \log_{10}$, based on the mean value reported by Strasburger (2001b) for the 10 afc task ($b = 9.67$ and $\beta = 11.44$; see Strasburger, 2001a for the conversions). Lapsing rate I is set to 2%. The functions are given by $p_c = g + (1-g-I) / (1 + 10^{b \lambda})$.

Stimulus parameters

The extent of the crowding effect depends upon the size of the stimuli, their separation distance, stimulus contrast, and the position in the visual field or in particular its retinal eccentricity (Bouma, 1970; Strasburger et al., 1991). So in order to study the effects at more than one retinal location with the same set of flanking distances while maintaining comparable perceptual conditions, stimulus sizes were M-scaled (equal sizes in the triplet), based on our earlier results (Strasburger et al., 1991, Strasburger, 2003a) and those of others. Table 1 gives the stimulus parameters. The scaling corresponds to an E_2 value of 2° visual angle in the equation

$$S = (1 + E/E_2) S_0 \tag{1}$$

(S : stimulus size, S_0 : foveal size, E : eccentricity; see Levi, Klein, & Aitsebaomo, 1985). Stimulus size was varied by changing the viewing distance (target height on the screen: 13.6 mm, 30×42 pixels; ring-cue diameter 20 mm). Distances were large enough that accommodation did not play a role.

Eccentricity (right meridian)	Size	Cue Diameter	Viewing Distance (cm)
1°	0.3°	0.44°	259
2°	0.4°	0.59°	194
4°	0.6°	0.88°	129.5

Table 1. Stimulus Parameters.

Procedure

The measurements were carried out by an experimenter who entered the subject responses on the computer keyboard. Subjects watched the stimulus monitor in a comfortable chair in a dimly lit room of constant low photopic illumination. A run of 30 trials (one threshold) took around 2 – 4 Minutes; short pauses were taken after ten runs or whenever the subject wished, to avoid fatigue. The design was within-subjects, all subjects ran all conditions. Data for a subject were typically collected in two sessions of $1\frac{1}{2}$ hours each. Thresholds in the three conditions were obtained in a balanced series to counter act series effects like learning and fatigue.

Results

Figure 3 shows the obtained mean recognition contrast thresholds over flanker distance. The thresholds for the single-character presentation against which the crowding conditions need to be compared are shown as thin horizontal lines. The standard error varied little between conditions so for clarity in the graphs it is only shown once in each subgraph, as error bars on the single condi-

tion. On average, the standard error was 0.069 log units or a factor of 1.17.

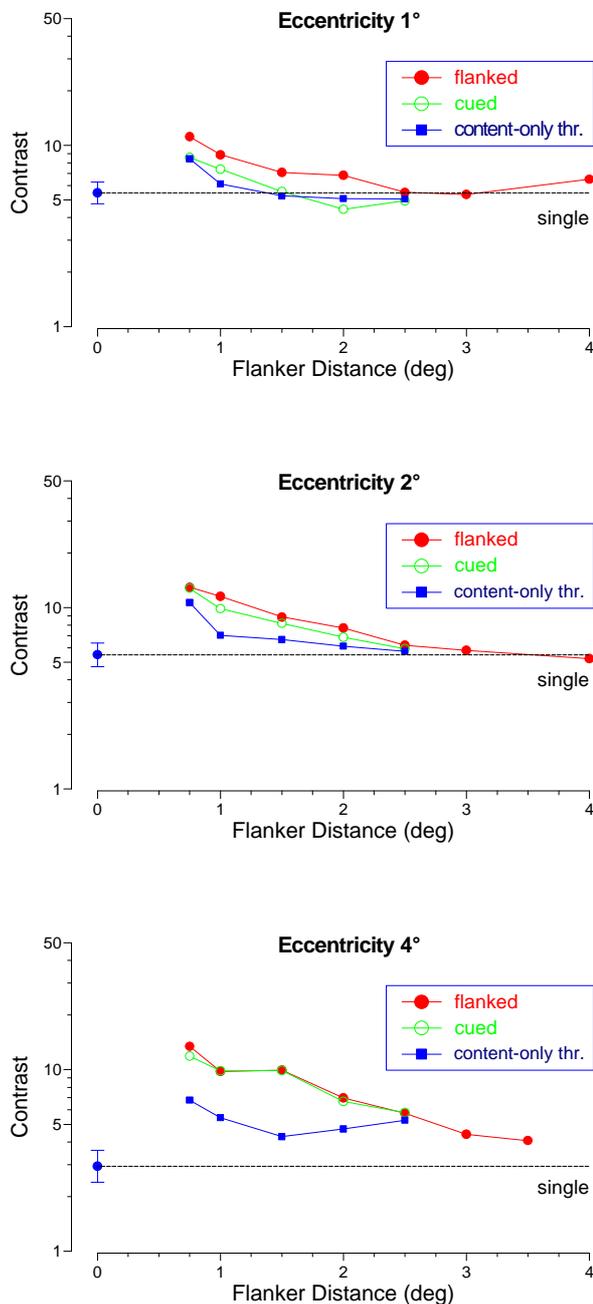


Figure 3. Recognition contrast thresholds for the three crowding conditions as a function of flanker distance, at three eccentricities on the horizontal meridian (top to bottom graph 1°, 2°, and 4°, respectively). The thresholds for the single-character presentation are shown as thin horizontal lines; the error bars on the corresponding data point show the mean for all the data points in that sub-graph.

As expected, at all three eccentricities the standard crowding condition (filled circles) gives the highest thresholds. Crowding is absent at sufficiently large flanker distances, as seen in the top and middle graph, and gradually and systematically sets in at decreasing distance. Obviously, Bouma's (1970) rule-of-thumb that crowding sets in at approximately a flanker distance of half the eccentricity value can only be a rough guide for selected conditions (Bouma specifies the flanker distance as free space, so his distance values are by one letter width smaller than those given here; see Discussion).

Contrast thresholds in the cued condition are always below that of the standard flanked condition (for significance see the ANCOVA results below) but far from that in the unflanked ("single") condition. Thus, except for the character size at 4° eccentricity, the cue was partially effective in focusing attention at the central character. Reasons for the cue not being more effective could be a too small circle size, such that some masking was introduced along with the capture of attention (Averbach & Coriell, 1961). Possibly, a steeper *M*-scaling function is needed for a positional cue than for other visual functions. Another effect could be a general decrease of the effects of positional cueing with increasing eccentricity. More on the cue's role will be apparent from the analysis of the correspondences given below.

Contrast thresholds are lowest in the condition of loosened attentional focus (filled squares). For 1° and 2° eccentricity – except for the closest character spacing – they are nearly equal to those in the single presentation (dashed line). At 4° eccentricity, these thresholds are elevated relative to the single condition but still far below those of the standard flanked condition. Thus, in a large part of the region where crowding occurs, the recognition of a character irrespective of where in a string it is, is nearly as good as that of a singular character.

Analysis of the Correspondences

Figure 4 shows the correspondences of the observers' incorrect responses with one of the flanking characters in the flanked condition, i.e. what Eriksen in his original flanker paradigm termed "error analysis" (C. W. Eriksen & Rohrbaugh, 1970; for a review of the later flanker task see C. W. Eriksen, 1995). A total of 4,105 errors were made in the flanked condition. Chance performance is indicated by the dashed line in the figure. It is at $1/9 + 1/8 = 23.6\%$, since the flankers are different from the central character and different from each other.

The difference between the proportion of correspondences and chance level can be interpreted as the proportion of localization errors. These are the responses where the observer correctly and not by chance reports a pattern content, but missed the location. As apparent from the figure, these localization errors are absent at sufficiently large flanker distances – at 3 degrees and more – but are increasingly evident with decreasing flanker distance, and

the more so the more peripheral the position of the central character is. At their maximum, the correspondences are as high as 52%, in other words there are close to 30% recognitions at the wrong location and only 48% true errors. Note that these performance figures refer to a situation where a majority of stimuli are at or around the contrast threshold, where of necessity (by the definition of threshold) failure of recognition occurs in part (in 45% of the cases).

The rising of correspondences above chance level with decreasing flanker distance is gradual, but nevertheless one can determine an approximate point at which it sets in. Table 2 shows the flanker distances where performance is still at chance (with respect both to the correspondences in Figure 4 and the contrast thresholds in the flanked condition in Figure 3). Even though these critical distances are only approximate, their similarity is apparent. Thus, localization errors occur if and only if there is crowding.

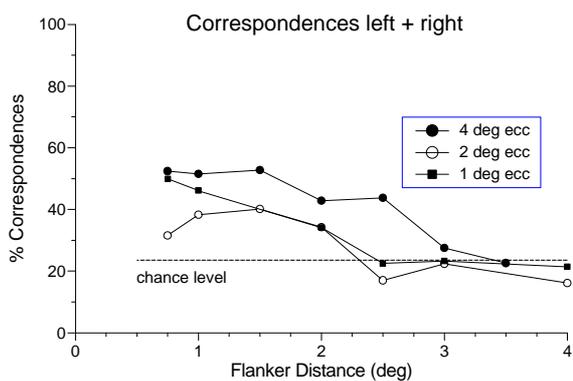


Figure 4. Correspondences of the observers' incorrect responses with one of the flanking characters in the flanked condition, as a function of flanker distance. Chance level (at 23.6%) is indicated by the dashed line.

Eccentricity	Critical distance	
	for correspondences	of thresholds (in flanked condition)
1°	2.5°	2.5°
2°	2.5°	2.5°
4°	3°	3°

Table 2. Approximate Critical Distances.

Results of the ANCOVA

To test the effect of the three crowding conditions on contrast thresholds and on correspondences for statistical significance, two analyses of covariance were performed, with crowding condition as a factor. Since the perceptual

performance in our paradigm strongly depends on retinal eccentricity and flanker distance (Strasburger et al., 1991), the variance introduced by the latter two variables was removed by introducing them as covariates. To linearize their influence (which improves the removal of variance), derived measures were used in the ANCOVA according to

$$M = 1 / (2^\circ + E) \quad (E: \text{eccentricity}) \quad (2)$$

$$d_i = 1 / d \quad (d: \text{flanker distance}) \quad (3)$$

Equation 2 linearizes the dependency on retinal eccentricity of many visual performance measures, as Weymouth (1958) has already shown. That Equation 3 will improve linearity can be seen by noting that the curves in Fig. 3 can be approximated by a hyperbola.

The target measure in the first ANCOVA was the increase $\Delta \log C$ of contrast threshold in one of the crowding conditions over that in the single condition, i.e.

$$\Delta \log C = \log C_c - \log C_s = \log C_c / C_s \quad (4)$$

with C_c : contrast in one of the three crowding conditions, and C_s : contrast in the single condition. The ANCOVA design was 3 conditions \times 5 flanker distances \times 3 eccentricities \times 20 subjects. Its results are summarized in Table 3.

Condition	Least Squares Means			
	log contrast increment	contrast incr. as factor	t Value (raw)	Pr > t (raw)
cued	0.2288	1.694	5.9***	< 0.0001
flanked	0.2808	1.909	7.25***	< 0.0001
content-o.	0.1304	1.350	3.36**	0.0018

condition	vs. condition	Differences of Least Squares Means			
		log contrast ratio	contrast ratio	t Value (raw)	Pr > t (raw)
cued	flanked	-0.05197	0.887	-3.23**	0.0026
cued	content-o.	0.09844	1.254	6.11***	< 0.0001
flanked	content-o.	0.1504	1.414	9.34***	< 0.0001

Table 3. ANCOVA of contrast thresholds. (DF = 38; significance levels, Bonferroni corrected: *5%, **1%, ***1%)

As the first part of the table shows, all three crowding conditions are effective in increasing the contrast threshold over that in unflanked presentation. The effect is highest for the flanked condition, with a mean increment, after the influences of retinal position and flanker distance are removed, of 0.28 log units (a factor of 1.9). The mean effect is smaller (0.23 log units) when a cue is added, and is smallest but still highly significant for the content-only threshold condition. Thus, even though in the latter the thresholds are much better than in a standard (flanked) crowding condition, there does remain a loss of sensitivity relative to a single-character presenta-

tion. It amounts to about half the log increment (0.13) than that of pure crowding (0.23).

The second part of the table shows that all three crowding conditions differ highly significantly from each other. Thus, in particular the difference between the cued and flanked condition – i.e. the influence of the cue – is large enough at 1° and 2° eccentricity to compensate for the rather small difference at 4° eccentricity, so that the latter can be safely interpreted as a peculiarity of the stimulus arrangement at 4° rather than as a general effect.

Table 4 shows the results for the ANCOVA of correspondences of observer responses with either one of the flankers. For the flanked and for the cued condition they are, on average (over flanking distances and contrasts), at 39%. (The somewhat lower rate of correspondences in the content-only threshold condition [34% still far above chance level of 23.6%] is an indirect effect of the fact that on average contrast is set lower by the adaptive algorithm. Since sensitivity is best in the center of attention, fewer characters outside that center are recognized.) The central point here is that these correspondences are all highly significantly above chance (1‰ level in all three cases). So there is – beyond doubt – recognition at locations *other* than the intended one.

Importantly, the correspondences are nearly equal between the flanked and the flanked-and-cued condition ($p=0.78$). As such, even though the cue is effective in improving recognition performance (as shown above), this improvement does not stem from removing attention from the flanking characters.

condition	Least Squares Means			
	estim. proportion of corresp's	Error	t Value (raw)	Pr > t (raw)
cued	0.3908	0.02971	13.15***	<.0001
flanked	0.3859	0.02971	12.99***	<.0001
content-o.	0.3353	0.02971	11.28***	<.0001

condition	Differences of Least Squares Means			
	vs. condition	Standard Estimate	t Value (raw)	Pr > t (raw)
cued	flanked	0.00487	0.28	0.7774
cued	content-only thr.	0.05552	3.25**	0.0024
flanked	content-only thr.	0.05066	2.96*	0.0052

Table 4. ANCOVA of Correspondences. (Significance level, Bonferroni corrected: *5%, **1%, ***1‰)

Table 5 shows the correspondences with either flanker broken up into the proportions for those with the left and the right flanker, separately. There seem a little more correspondences with the left than with the right flanker. Since, however, the difference is small and the stimulus presentation was not balanced with respect to laterality (presentation was on the right meridian only) this difference is not further considered.

	Correspondences (%)		
	left	right	either
flanked	20.2	18.4	38.6
cued	21.8	17.3	39.1
loose att.	16.0	17.5	33.5

Table 5. Correspondences with left and right flankers.

Discussion and conclusions

Three effects are striking and need to be discussed: (1) At intermediate flanker distances, from ~ 3° down to 1° and less center-to-center distance, the crowding effect is to a large part explained by imprecise coding of the character to be recognized (i.e. is for a major part reduced when positional errors are treated separately). (2) A ring cue, presented in advance of the target enhances recognition performance and in this sense sharpens spatial attention, but at the same time leaves the positional coding equally imprecise. (3) Pattern content and pattern location are coded separately.

Gestalt and closedness

With respect to effect (1), the results lend new support to the Gestalt idea that the whole is more than the sum of its parts: Whatever the components or the features of any one of the three characters are (Gabor patterns, wavelets, edges, T junctions, end points), they are processed together, such that any one (or more) of the characters are recognized, somewhat irrespective of their location. The whole (the character) is not just the sum of its parts (features) but includes some mechanism of binding those together. The joint processing of features is induced by the characters being separated by a gap, or, as Gestalt theory would say, by the closedness of their Gestalt. The uncertainty which character is to be processed increases with decreasing distance (as is seen from the localization errors; Fig. 4), but the processing itself remains much less affected, as is apparent from the content-only thresholds (in which the locus of processing is left free) being similar to the single-character thresholds. It seems as if an attentional spotlight can focus on any one of the characters, with a preference at the intended location. This focusing-in becomes less precise with increasing eccentricity in the visual field. An impreciseness of coding the letter position has already been shown by Eriksen & Rohrbaugh (1970) for a circular stimulus arrangement of 2.2° radius. Recently, Huckauf & Heller (2002b) have shown that the impreciseness increases with eccentricity (1° – 7°). Chung et al. (2003) have provided additional evidence for the same, using a linear arrangement. They find that the standard deviation of position noise (as they call it) scales with letter size and increases with eccentricity, from 0.5 letter spaces in the fovea to 0.96 letter spaces at 4 letter positions from fixation.

Role of the ring cue

The role of the ring cue is intriguing. As the analysis of variance has shown, the ring cue does improve recognition performance. In common understanding this is evidence for spatial attention being concentrated at the target, either by enhancing neural activity at the target position or by suppressing activity at neighboring positions. The cue is, however, not effective in reducing correspondences with flanking characters and is, in this sense, not effective in removing attention from the flankers. Thus, the ring cue can be said to increase attention at the target without reducing attention at surrounding locations.

This result might be understood in terms of the distinction between transient and sustained attention (Nakayama & Mackeben, 1989; Mackeben, 1999; see the section on feature integration below). The ring cue, appearing shortly before the target, would be expected to attract transient attention, whereas the instruction to the subject to attend to the center character is expected to mobilize sustained attention. Note that this account would imply that transient attention does not improve the positional coding in indirect or peripheral vision.

For the performance enhancing effect, the timing and the size of the ring cue is important as is known from previous work. The timing seems well chosen in the present experiments as evidenced by the results of the ANCOVA. The cue size was perhaps a little too small and (from the results at 4° eccentricity) might need to be scaled with eccentricity at a steeper rate than the increase of receptive-field diameter.

Separate coding of pattern content and location?

The difference between the proportion of correspondences and chance level has been interpreted as the proportion of localization errors above. What underlies these responses? As an anonymous reviewer has put it, are these “true” mislocalization errors in the sense that the observer perceives the identity of the target to be that of the flanker? The latter would imply that the observer was unaware of the difference in identity, i.e. it raises the question of how conscious the mislocalization was. Subjects when asked, are mostly unaware of the error. Awareness might be a separate process, though, and a more neutral way to state the assertion is that pattern content is, at a certain stage, coded independent of its location.

Huckauf & Heller (2002a, p.122), (who find that flanker onset is of critical importance) speak of failures in “the process of coding a letter as being at a target location”. An alternative interpretation, offered by the same reviewer, is, that observers are “just biased to respond in that manner when they are unable to identify the target”. In other words, a pattern-recognition mechanism in the brain has extracted the flanker’s pattern content and that information then influences the response system when

the sensory input from the target is weak. Still, however, the pattern content is treated somewhat irrespective of its location, perhaps at a later processing stage. Based on our paradigm, we cannot decide between such possibilities. What we can conclude is that the pattern content at some stage in processing acts independent of the pattern’s location. Separate coding of pattern content and pattern location at that stage seems a simple way to summarize the basis of the phenomenon.

Flanker distance: Bouma’s rule

Bouma (1970) has formulated a memorable rule-of-thumb for the critical flanker distance needed in the discussion below which states that this distance (in the radial direction; Toet & Levi, 1992) amounts to approximately half the eccentricity value. Since Bouma’s rule is now more widely cited – Pelli, Palomares, & Majaj, 2004, Table 4, and Chung, Levi, & Legge, 2001, Table 1, both provide tables of results from the literature on that matter – it is useful to note that Bouma has defined the flanker distance differently than seems now more popular, namely as the gap between the target and the flanker, rather than the distance from center to center (as in this report and others, e.g. Strasburger et al., 1991, Chung et al., 2001, Pelli et al., 2004). If w denotes the width of the target stimulus, Bouma’s rule states

$$d_{\text{Bouma}} = 0.5 E = d - w \quad (5)$$

(d : center-to-center flanker distance, E : eccentricity, w : target width; all in degrees of visual angle). Bouma used character widths between 0.16° and 0.29° and eccentricities from 2° to 8°; so his rule would correspond to approximately

$$d \approx 0.5 E + 0.225^\circ \approx 0.57 E \quad (6)$$

The critical distance according to Bouma (1970) is, thus, somewhat *larger* than half the eccentricity value, and is around 0.6. Since it is not sharply defined, the precise form of that equation (linear, either passing through w or through the origin) is not important. Results by Pelli et al. (2004, Fig. 3b) suggest that the curve $d(E)$ is slightly nonlinear and curved upwards. Interestingly, Pelli finds typical slope values in the above equation to be somewhat *lower*, at 0.3 to 0.4.

Note that Bouma’s rule in its correct form as in Equation 5 is formally equivalent to M scaling, as in Equation 1 above, specifying that an item of spatial extent varies with retinal eccentricity by a linear function which does not pass through the origin:

$$d = 0.5 E + d_0 \quad (7)$$

with an E_2 value of $E_2 = w/0.5$ and a foveal d value of $d_0 = w$. Bouma’s rule (in the range that he studied), corresponds to an E_2 of about 0.45°. Chung et al. (2001), for example, find a d_0 of 0.5° and E_2 of 1.7°.

Theoretical accounts: focusing attention and the field of feature integration

In a comprehensive analysis, Pelli and coworkers (2004) have characterized crowding as a process of impaired feature integration occurring in the visual periphery, in contradistinction to (lateral) masking as occurring from impaired feature detection anywhere in the visual field. We have ourselves characterized the visual periphery – where the interesting cases of crowding occur (Strasburger et al., 1991) – as differing from the fovea by the architecture of feature integration (Strasburger & Rentschler, 1996). That argument was based on the differing dependence-on-eccentricity functions of contrast sensitivity for grating detection and for character recognition (Strasburger, Rentschler, & Harvey, 1994; Strasburger, Gothe, & Lutz, 2000; Strasburger, 2003b), and by showing that the difference between the two cannot be explained by a spatial scaling concept (M scaling, cortical-magnification scaling). We concluded that there must be architectural differences across the visual field – in particular between the fovea and the rest of the field – that concern feature integration, not feature detection. In a hierarchy of task complexity ranging from (1) pattern detection (present/non-present), (2) coarse grating discrimination¹ (horizontal/vertical), (3) fine grating discrimination (orientation threshold), and (4) character recognition or identification, tasks (1) and (2) can be treated as more or less equivalent and as different from (4). The distinctive characteristic of recognition (or identification) we considered to be the requirement for the observer to chose the responses from a comparatively large number of alternatives.

Pelli et al. (2004) have chosen Bouma's rule – that the critical distance scales with eccentricity – together with our finding (Strasburger et al., 1991) that this rule holds independent of signal size, as the defining criterion for crowding (as opposed to lateral masking). The area within that critical distance (we might call it Bouma's area) is interpreted by Pelli as the area of feature integration. What I propose here is that this field of feature integration is related to spatial attention and indeed could be the same as what in the attention literature is referred to as the attentional spotlight, and in light of the distinction of sustained and transient visual attention (Nakayama & Mackeben, 1989; Mackeben, 1999) could be called a *field of sustained visual attention*². The evidence is in the meaning of the word attention: subjects in our task were instructed to attend to the central character; they did so as well as they could and when asked reported that they needed to attend outside the point of fixation. Regarding the distinction between sustained and transient visual attention, our task classifies as the former since subjects were well aware in advance of where the stimulus would appear (within any run of about 30 trials the stimulus appeared always at the same location). What we cannot clarify based on our paradigm is whether that area (that

focus, spotlight) of attention includes all *three* characters. In other words, is there a separate mechanism responsible for choosing the central character and a further mechanism that defines the three areas of integration, or does the focus of attention zoom-in on the area of one of the three characters but is imprecise in doing so, both with respect to locus and size? Parsimony leads us to assume the latter, that all this is accomplished by a single, sustained visual-attentional mechanism.

The notion of features being processed together does not contradict the opposite, that of free-floating features (from the flankers) getting jumbled or miscombined. Tjan & Chung (2004), e.g., show the importance of similar spatial frequency between target and flanker which supports the latter. Indeed and correspondingly, the content-only thresholds presented here are still elevated relative to the single condition. (Graphically, feature binding relates to the area above the content-only thresholds in Fig. 3, and feature jumbling to the area below). Pelli et al. (2004) consequently assume that the area of feature integration in the crowding situation includes the target plus flankers, (so that features get jumbled across the characters). With our finding of feature binding, Pelli's assumption of a large integration field would thus require a further, smaller integration field contained within, or the presence of an attentional gradient, to integrate those features that belong to one but not another character. Our finding would then also contradict Pelli's idea that the periphery is lacking small enough integration fields (Theory point *a*, in Table 3, p. 1153). On the other hand, the interpretation favored here, namely a smaller integration field centered on the recognized character, but with a soft outline such that features from farther away get included, would encompass all the empirical findings discussed here.

The role of the positional cue is then one of modulating transient spatial attention. Thus, we assume further that there is a separate area of transient attention, often but perhaps not necessarily within the focus of sustained attention, that enhances either feature detection or feature integration.

For tasks that do not require feature integration to be solved, the characteristics as discussed above need not hold. This applies in particular to the now often-used task introduced by Flom et al. (1963) in their seminal paper on contour interaction, where the orientation of a Landolt-C embedded in bars is to be selected. Thus, the interpretation of the area of feature integration being the same as the sustained attentional spotlight forwarded here is consistent with Nazir's (1992) finding that precueing has no effect on crowding in the Flom task, as well as the report of Wilkinson et al. (1997) of the same for grating targets, since neither of the tasks involves feature integration. Similarly, the contradiction between Hess et al. (2000) and Kooi et al. (1994; see the paragraph on flanker similarity below), where Hess et al. find that in the near periphery flanker polarity plays no role in lateral masking and Kooi et al. (1994) and others find the opposite, can

be resolved by noting that Hess et al. (2000) used the Flom task that does not require feature integration and is thus not directly comparable to a character crowding task.

Huckauf et al. (1999) show limitations of the feature interaction account for explaining the crowding effect, but indeed their findings support both Pelli's (2004) conclusions and those presented here. Huckauf used trigrams of constant-size (0.51°) letters, rotated letters, and pseudo-letters of high contrast presented at eccentricities up to 7° and flanker distances up to 2° , and measured the percentage of correct recognition of the center stimulus. She shows that feature interaction cannot account for the differences between the crowding effects from different flanker types, and concludes that a higher-level mechanism ("letter level") above the feature detection stage must be of influence. That stage processes different flanker types differently and thus influences the amount of the crowding effect. According to Huckauf et al. (1999), feature jumbling is not the source of crowding. Note, however, that their paradigm directly tests the effect of flanker *type* only, which accounts for only 5% of the crowding effect's total variance. Whether the main effects of crowding – i.e. the effect of flanker presence, modulated predominantly by retinal eccentricity and flanker distance – are, or are not, mediated by feature jumbling, cannot be decided by their paradigm. So their claim is that of the participation of higher-level mechanisms, not of their exclusive effectivity.

The similarity of the flankers with the target plays an important role for the strength of lateral interaction (B. A. Eriksen & Eriksen, 1974; Nazir, 1992; Chung et al., 2001; Kooi et al., 1994). Kooi et al. (1994) show in particular that the crowding effect – for a letter T embedded in Ts – is radically reduced when the flankers are of the opposite contrast polarity. The authors discuss that finding with respect to the salience of the target within the flankers ("pop-out") and conclude (p. 276) that (cortical) mechanisms at a lower level than the attentional are likely to mediate crowding. I here propose the alternative explanation that the salience of the target (or the flankers) attracts a field of feature integration – a field of spatial attention – and if the salience is weaker, then recognition is not impaired but is indeed taking place unimpaired, just at the "wrong" location. The position code is lost, as is the information from which eye the stimulus was received (Kooi et al., 1994). Similarity and salience thus govern *where* recognition takes place, not *how*.

The concept of an attentional spotlight dates back at least to Johannes Müller in the early 19th century and was empirically supported already by von Helmholtz (1871; see A. H. C. van der Heijden, 1992). It has received widespread attention in the last decade both in the behavioral (e.g. Tsai & Lavie, 1988) and biological cognitive sciences, and it is now widely agreed upon that with a few amendments (e.g. Stoffer, 1988) the spotlight metaphor, with a spotlight of variable size and locus, is a concise, empirically well supported concept capturing many aspects of

visual spatial attention. (Stoffer, 1988 proposes to model spatial attention not by single, movable spotlight but by a battery of spotlights of various sizes and loci, that can get turned on and off.) In Bundesen's related computational Theory of Visual Attention (TVA) (Bundesen, 1990, 1998), two selection processes are postulated, one that prioritizes objects ("pertinence"), and one that prioritizes features ("pigeonholing"). Within that framework, the results presented here would mean that the pertinence selection mechanism often selects the wrong character. In summary I propose here that an attentional spotlight, with a focus size larger than one of the characters, is *voluntarily but only approximately* pointed at a locus in the visual field where the trigram of characters is, and further, that within that field, a field of feature integration is established, the latter sharpened in its boundaries by pattern segregation algorithms (Gestalt formation). A transient occurring at or near (or pointing towards) the actual locus of attention, like the suddenly occurring ring cue here, will further enhance processing in the attentional field, but *without* changing its locus.

Flanker contrast

Chung et al. (2001, Fig. 9) and Pelli et al. (2004, Fig. 10) used a crowding stimulus arrangement similar to the one introduced by us and used here (trigrams of characters of same size and contrast where the recognition contrast threshold of the center character was the criterion variable; Strasburger et al., 1991), but used different contrast between target and flanker to study the dependence of crowding on the contrast of the flankers. Both studies report that the flankers become effective in masking only when they have the same as, or higher contrast than, the target, with a monotonic increase (i.e. no facilitatory effects) from that point on. Pelli et al. (2004, Fig. 10 a & b) report that the corresponding function of contrast threshold for target recognition versus flanker contrast has a peculiarly high slope and fast saturation, similar to a psychometric function. Furthermore, they report that the critical flanker contrast for crowding (the lower knee of that function) is independent of the flanker distance. Even though both groups of authors frame their results in a bottom-up theoretical account, these findings fit rather well with the ideas presented here: A steep slope of the psychometric function is, as we have shown previously (Strasburger, 2001b), characteristic for target recognition or identification as opposed to target detection or discrimination. Thus, Chung's and in particular Pelli's findings are accounted for by assuming that their flanker-effect function reflects the psychometric function for the *recognition of the flankers*. The critical flanker contrast in Pelli et al. (2004) would then be the recognition contrast threshold of the flankers. This would predict all three, the steep slope, saturation, and the independence of the critical flanker contrast of flanker distance. The flanking characters will sufficiently attract feature integration fields

(sustained attention fields) if and only if their contrast is comparable to the alternative patterns in the field, and, if that is the case, character identification can take place in the field of the flanker, provided the integration field is inadvertently centered there. The notion that crowding occurs through unfocused attention is thus compatible with that of information being pooled over the extent of the trigram, as Chung et al. (2001) and Pelli et al. (2004) conclude, but it also goes beyond that interpretation.

Letters and digits

Note that the present results were obtained with trigrams of digits and will with respect to semantic influences not directly generalize to letter trigrams: There is a word-superiority effect for letter strings such that the center character is more easily recognized when it occurs within a word than in a meaningless string (e.g. Fine, 2001). In digit strings this cannot happen since all strings are equally meaningful. In a sense, digits are more suitable than letters to test characteristics of pure pattern recognition and its underlying mechanisms since the semantics play no differential role.

Care should be taken when the standard crowding task (recognizing the center character) is compared with a task where the surround is also attended to as in recognizing a three- (or more) letter word. The attentional focus can change in size and scale (e.g. Stoffer, 1994) thereby changing its characteristics, as the research on Navon patterns (Navon, 1977) has shown. As an aside, note that reading usually takes place at high contrast and comparatively small letter size so that for the involved pattern recognition process a predominant mediation by the parvo system is commonly assumed (e.g. Vidyasagar, 2001), whereas at the contrast threshold of character recognition a dominant or even exclusive involvement of the magno system is rather likely (Strasburger & Rentschler, 1996).

Neural mechanisms of crowding

How does the proposed model fit in with current theorizing on brain mechanisms of crowding? Flom et al. (1963) have shown that the lateral interaction in their paradigm occurs also in dichoptic presentation – when the target and flanker are presented to different eyes. They conclude that the neural basis of the interactions must be at a stage where the information from both eyes has come together, i.e. must be cortical. Several researchers have since verified the finding of dichoptic masking, in the fovea and in the periphery. A particularly striking example is presented by Tripathy and Levi (1994) who show dichoptic masking at the blind spot, with the flankers around the blind spot and the target presented to the other eye at the blind spot's corresponding position. This is amazing since the blind spot's projection area in the primary visual cortex – i.e. in that part of the visual cortex that is most distinctly organized retinotopically – is believed to be monocular (e.g. LeVay, Connolly, Houde, &

Van Essen, 1985). Tripathy and Levi (1994) propose this is solved by long-range horizontal connections in the striate cortex that have not been identified anatomically yet.

Our results and theoretical account allow for a simpler neural interpretation (that does not require such horizontal connections). The concept of spatial attention relies on a spatially selective enhancement or inhibition of neural bottom-up activation through top-down mechanisms. The selectivity is assumed to be mediated by retinotopically organized brain structures – referred to as maps – that need not be cortical (see LaBerge, 1995, and Vidyasagar, 2001, for reviews; suitable candidate structures are the lateral and inferior pulvinar of the thalamus [LaBerge, 1995; LaBerge & Brown, 1989] or the reticular nucleus of the thalamus [Crick, 1984, based on Skinner & Yingling, 1977]). The gating itself could happen in early cortical visual areas or even in the lateral geniculate nucleus, which is commonly assumed to subserve a gating function in the retino-cortical pathway. For example, Vidyasagar (1998, 2001) has shown attentional modulation in single-cell studies as early as in V1. In his 2001 model, he proposes that the attentional spotlight receives its steering information from the magno pathway which would explain why the low-contrast characters used here can steer the “spotlight” (Strasburger & Rentschler, 1996).

As a working hypothesis, then, frontal and posterior-parietal regions involved in attention could – mediated through the pulvinar and V1 – selectively enhance or suppress retinotopically organized bottom-up activation such that, by a subsequent winner-take-all network, the dominant stimulus representation is the only one processed in the higher-up areas involved in visual feature integration (like the fusiform gyrus in humans or the inferotemporal cortex in the macaque; Fujita, Tanaka, Ito, & Cheng, 1992; Tanaka, 1996). Hence, if the neural map information is imprecise in locus or spatial extent, the feature integration area will operate in an unintended visual field area, so that the perceived pattern does not coincide with the target. The corresponding observer response will be counted as incorrect, which leads to a decreased proportion of correct responses for the target and a seemingly increased contrast threshold. The pre-appearing ring cue, on the other hand, is assumed to pre-activate the corresponding (retinotopic) location in the map without influencing other locations. Note that any such attentional “top-down” explanation as proposed here will not contradict but rather complement low-level masking accounts favored by others where masking might already take place in the retina.

Where and what

In a surprising way, our results lend psychophysical support to the now widely held proposition of separate neural processing – and independent coding – of pattern content and pattern location, i.e. of the *what* and *where*

being processed in a ventral and a dorsal stream, respectively (Ungerleider & Mishkin, 1982; Ungerleider & Haxby, 1994). According to our results, when an object is recognized in indirect vision, more often than not the position code is imprecise or is lost, so that a pattern may be recognized, while it is not clear which pattern that was. Concisely put, the fusiform gyrus may sometimes correctly integrate pattern features while the posterior parietal cortex cannot decide unambiguously where attention to that pattern was focused at.

Acknowledgements

My thanks go to Dorothe Poggel for valuable discussions on attention and neuropsychology and her dedication and supportiveness in my group, to Ingo Rentschler for continuing support and funds, Ernst Pöppel for his enthusiasm and support at the GRP, Dirk Vorberg for pointing out early on the importance of manipulating stimulus variables believed to influence spatial attention, to Manfred MacKeben for helpful comments on the manuscript, to David Carr who did the analysis of covariance, Ilona Wolff for running the subjects in the main experiment, Rainer Wolff for doing early statistical analyses, Susanne Hörz for data entry, and Andreas Bohne, Elke Berger, Ulrike Bunzenthall and Dorothe Poggel for being unpaid subjects in the pilot experiments in Magdeburg.

Part of the study was funded by the Deutsche Forschungsgemeinschaft, grant STR 354/3-1 to the author. A preliminary report was presented at the Cognitive Neuroscience Society Meeting 2002 (Strasburger & Poggel, 2002) and the VSS 2003.

Commercial relationships: none.

Email: strasburger@med.uni-goettingen.de;
strasburger@uni-muenchen.de

Address: Dept. Med. Psychology, University of Göttingen, Waldweg 37, 37073 Göttingen, Germany.

Footnotes

¹ The term “discrimination task” is sometimes used in a different meaning, implying the judgement of a quantity being larger or smaller than another (the corresponding psychometric function then goes from -1 to 1). This is not implied here, the intended meaning being that the observer can discriminate between two broadly different stimuli and thereby identify each. The term “identification task” is sometimes used for that case but is avoided here to reserve the concept of identification for those tasks where discrimination between a few cases will not solve the identification.

² The term “sustained attention” is used in a different, unrelated meaning in the more general attention literature, referring to alertness and vigilance.

References

- Atkinson, J., Pimm-Smith, E., Evans, C., Harding, G., & Braddick, O. (1986). Visual crowding in young children. *Documenta Ophthalmologica Proceedings Series*, 45, 201-213.
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *The Bell System Technical Journal*, 40, 309-328.
- Bach, M., Meigen, T., & Strasburger, H. (1997). Raster scan cathode ray tubes for vision research - limits of resolution in space, time and intensity, and some solutions. *Spatial Vision*, 10(4), 403-414. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177-178. [PubMed]
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523-547. [PubMed]
- Bundesen, C. (1998). Visual selective attention: Outlines of a choice model, a race model, and a computational theory. *Visual Cognition*, 5, 287-309. [Abstract]
- Chung, S. T. L., Legge, G. E., & Ortiz, A. (2003). Precision of local signs in central and peripheral vision. *Journal of Vision*, 3(9), 815a. [Abstract]
- Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research*, 41(14), 1833-1850. [PubMed]
- Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences, USA*, 81(14), 4586-4590. [PubMed] [Article]
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of visual attention. *Annual Reviews of Neuroscience*, 18, 193-222. [PubMed]
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143-149.
- Eriksen, C. W. (1995). The flanker task and response competition: a useful tool for investigating a variety of cognitive problems. *Visual Cognition*, 2, 101-118.
- Eriksen, C. W., & Collins, J. F. (1969). Temporal course of selective attention. *Journal of Experimental Psychology*, 80(2), 254-261. [PubMed]
- Eriksen, C. W., & Rohrbaugh, J. W. (1970). Some factors determining efficiency of selective attention. *American Journal of Psychology*, 83, 330-343.

- Fine, E. M. (2001). Does meaning matter? The impact of word knowledge on lateral masking. *Optometry and Vision Science*, 78(11), 831-838. [PubMed]
- Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effects. *Science*, 142, 979-980. [PubMed]
- Flom, M. C., Weymouth, F. W., & Kahnemann, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America*, 53(9), 1026-1032. [PubMed]
- Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360, 343-346. [PubMed]
- Gazzaniga, M. S. (Ed.). (1999). *The New Cognitive Neurosciences* (2nd ed.). Cambridge, MA: Bradford Book, MIT Press.
- Geiger, G., & Lettvin, J. Y. (1989). Dyslexia and reading as examples of alternative visual strategies. In C. von Euler (Ed.), *Brain and Reading* (pp. 331-343). London: Macmillan.
- Harvey, L. O., Jr. (1986). Efficient estimation of sensory thresholds. *Behavior Research Methods, Instruments, & Computers*, 18(6), 623-632. [Article]
- Harvey, L. O., Jr. (1997). Efficient estimation of sensory thresholds with ML-PEST. *Spatial Vision*, 11(1), 121-128. [PubMed]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383, 334-337. [PubMed]
- Hess, R. F., Dakin, S. C., Kapoor, N., & Tewfik, M. (2000). Contour interaction in fovea and periphery. *Journal of the Optical Society of America A*, 17(9), 1516-1524. [PubMed]
- Huckauf, A., & Heller, D. (2002a). Spatial selection in peripheral letter recognition: In search of boundary conditions. *Acta Psychologica*, 111(1), 101-123. [Abstract]
- Huckauf, A., & Heller, D. (2002b). What various kinds of errors tell us about lateral masking effects. *Visual Cognition*, 9(7), 889-910. [Abstract]
- Huckauf, A., Heller, D., & Nazir, T. A. (1999). Lateral masking: limitations of the feature interaction account. *Perception & Psychophysics*, 61(1), 177-189. [PubMed]
- Jüttner, M., & Strasburger, H. (1997). FORPXL - A Fortran interface to PXL, the psychological experiments library. *Spatial Vision*, 10(4), 491-493. [PubMed]
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, 8(2), 255-279. [PubMed]
- LaBerge, D. (1995). Computational and anatomical models of selective attention in object identification. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (1st ed.). Cambridge, MA: MIT Press.
- LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, 96(1), 101-124. [Abstract]
- LeVay, S., Connolly, M., Houde, J., & Van Essen, D. (1985). The complete pattern of ocular dominance stripes in the striate cortex and visual field of the macaque monkey. *Journal of Neuroscience*, 5(2), 486-501. [PubMed] [Article]
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25, 963-977. [PubMed]
- Mackeben, M. (1999). Sustained focal attention and peripheral letter recognition. *Spatial Vision*, 12(1), 51-72. [PubMed]
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29(11), 1631-1647. [PubMed]
- Navon, D. (1977). Forest before trees. *Cognitive Psychology*, 9, 353-383. [Abstract]
- Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research*, 32(4), 771-777. [PubMed]
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12), 1136-1169. [PubMed] [Article]
- Schneider, W. X., & Mojzisch, A. (2001). Visuelle Aufmerksamkeit und Kognitive Neurowissenschaft: Befunde und Mechanismen. (Visual attention and cognitive neuroscience: Findings and mechanisms). In M. Eimer & T. Goschke (Eds.), *Enzyklopädie der Psychologie: Band Kognitive Neurowissenschaft*. Göttingen: Hogrefe.
- Skinner, J. E., & Yingling, C. D. (1977). Central gating mechanisms that regulate event-related potentials and behavior. In J. E. Desmedt (Ed.), *Progress in clinical neurophysiology: Attention, voluntary contraction and event-related cerebral potentials, Vol. 1* (pp. 30-69). Basel: Karger.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74(11).

- Stoffer, T. H. (1988). *Dynamische Aspekte der visuellen Aufmerksamkeit. Funktionelle Charakteristika der Fokussieränderung vom Typ "Gummilense" und ihre Beteiligung an der Entstehung der Dominanz globaler über lokale Merkmale*. Habilitationsschrift, Universität Bielefeld.
- Stoffer, T. H. (1994). Attentional zooming and the global-dominance phenomenon: effects of level-specific cueing and abrupt visual onset. *Psychological Research*, 56(2), 83-98. [PubMed]
- Strasburger, H. (1997). *R_Contrast*: Rapid measurement of recognition contrast thresholds. *Spatial Vision*, 10(4), 495-498. [PubMed]
- Strasburger, H. (2001a). Converting between measures of slope of the psychometric function. *Perception & Psychophysics*, 63(8 (special issue *Psychometric Functions and Adaptive Methods*)), 1348-1355. [PubMed]
- Strasburger, H. (2001b). Invariance of the psychometric function for letter recognition across the visual field. *Perception & Psychophysics*, 63(8 (special issue *Psychometric Functions and Adaptive Methods*)), 1356-1376. [PubMed]
- Strasburger, H. (2003a). A generalized cortical magnification rule predicts low-contrast letter recognition in the visual field. *Journal of Vision*, 3(9), 653a. [Abstract]
- Strasburger, H. (2003b). *Indirektes Sehen. Formerkennung im zentralen und peripheren Gesichtsfeld*. Göttingen, Bern, Toronto, Seattle: Hogrefe. [Abstract]
- Strasburger, H., Gothe, J., & Lutz, K. (2000). The healthy visual field of recognition. *Perception*, 29 Suppl., 84-85. [Abstract]
- Strasburger, H., Harvey, L. O. J., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and excentric view. *Perception & Psychophysics*, 49, 495-508. [PubMed]
- Strasburger, H., & Poggel, D. A. (2002). *The visual crowding effect is caused, in part, by unfocused spatial attention*. Paper presented at the Cognitive Neuroscience Society (CNS 2002), San Francisco.
- Strasburger, H., & Rentschler, I. (1995). Is the crowding effect of purely attentional origin? *Perception*, 24, Suppl., 77. [Abstract]
- Strasburger, H., & Rentschler, I. (1996). Contrast-dependent dissociation of visual recognition and detection field. *European Journal of Neuroscience*, 8, 1787-1791. [PubMed]
- Strasburger, H., Rentschler, I., & Harvey, L. O., Jr. (1994). Cortical magnification theory fails to predict visual recognition. *European Journal of Neuroscience*, 6, 1583-1588. [PubMed]
- Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology*, 53, 471-477. [PubMed]
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19, 109-139. [Article]
- Tjan, B., & Chung, S. T. L. (2004). Form vision in the periphery. *Journal of Vision*, 4(11), 88a. [PubMed]
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32(7), 1349-1357. [PubMed]
- Tripathy, S. P., & Levi, D. M. (1994). Long-range dichoptic interactions in the human visual cortex in the region corresponding to the blind spot. *Vision Research*, 34(9), 1127-1138. [PubMed]
- Tsal, Y., & Lavie, N. (1988). Attending to color and shape: the special role of location in selective visual processing. *Perception & Psychophysics*, 44(1), 15-21. [PubMed]
- Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current opinion in neurobiology*, 4, 157-165. [PubMed]
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.
- van der Heijden, A. H. C. (1987). Central selection in vision. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 421-446). Hillsdale, N. J.: Erlbaum.
- van der Heijden, A. H. C. (1992). *Selective attention in vision*. London, New York: Routledge.
- Vidyasagar, T. R. (1998). Gating of neuronal responses in macaque primary visual cortex by an attentional spotlight. *NeuroReport*, 9, 1947-1952. [PubMed]
- Vidyasagar, T. R. (2001). From attentional gating in macaque primary visual cortex to dyslexia in humans. *Progress in Brain Research*, 134, 297-312. [PubMed]
- von Helmholtz, H. (1871). Über die Zeit, welche nötig ist, damit ein Gesichtseindruck zum Bewusstsein kommt. *Berliner Monatsberichte, Juni*, 333-337.
- Weymouth, F. W. (1958). Visual sensory units and the minimal angle of resolution. *American Journal of Ophthalmology*, 46, 102-113. [PubMed]
- Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A*, 14(9), 2057-2068. [PubMed]